Corruption, Intrinsic Motivation and the Love of Praise

Amrita Dhillon, Antonio Nicolò and Fei Xu*

October 10, 2017

*Department of Political Economy, King's College London, (amrita.dhillon@kcl.ac.uk); Antonio Nicolò, Department of Economics and Management, Università degli Studi di Padova and School of Social Science, University of Manchester, Arthur Lewis Building-3.018, Manchester, M13 9PL (antonio.nicolo@unipd.it); Fei Xu, School of Social Science, University of Manchester, (email: fei.xu@manchester.ac.uk). We thank Aislinn Bohren, Rabah Amir, David Dillenberger, Hanming Fang, Sangmok Lee, Andrew Postlewaite, and Marie Claire Villeval for valuable suggestions. Fei is grateful to the Department of Economics at the University of Pennsylvania for the kind hospitality during 2016 academic year.

Abstract: Do higher wages prevent corruption (bribe taking)? We investigate a setting where individuals who apply for public sector jobs are motivated not just by monetary incentives but also by intrinsic motivation and concern for the collective reputation of their profession. We show that an increase in monetary compensation may cause reputation concerned individuals to be more prone to participate in corruption due to an "overjustification" effect. The overall effect of monetary incentives on fighting corruption crucially depends on the composition of the pool of public sector workers for two reasons: first, different types of workers react differently to the same policy; second, the composition of the pool of workers affects individual behaviour through its effect on collective reputation. These results imply in particular that policies to fight corruption should focus more on increasing the collective reputation of the public sector rather than using monetary incentives which have perverse effects on some agents.

Keywords: Corruption, Collective Reputation, Intrinsic and Extrinsic Motivation.

JEL Classification Number: A13, D73, H10, Z13

1 Introduction

" *Corresponding to the three types in the city, the soul also is tripartite, [...] we speak of this [the first] part of the soul, and justify our calling it the money-loving and gain-loving part? And [the second part] it is wholly set on predominance and victory and good repute [...] But surely it is obvious to everyone that all the endeavour of the [third] part by which we learn is ever towards knowledge of the truth of things, and that it least of the three is concerned for wealth and reputation [...] And that is why we say that the primary classes of men also are three, the philosopher or lover of wisdom, the lover of victory and the lover of gain."*

—Plato in the Republic book 9, sections 580d-581c

Low wages are often cited as a cause of bureaucratic corruption (see Becker and Stigler, 1974; Chand and Moene, 1997; Amir and Burr, 2015) and the use of "efficiency" wages to fight bribery has been proposed as a solution (see Rijckeghem and Weder, 2001; Mahmood, 2005). However, evidence for the effect of monetary incentives as an anti-corruption measure is still quite mixed (see e.g., Treisman, 2000 and Swamy et al. 2001). A large body of field and experimental evidence indicates that monetary incentives and punishment may actually serve as "negative reinforcers" for the desired behaviour because they sometimes conflict with non-pecuniary motivation (see, e.g., Titmuss, 1970; Akerlof and Dickens, 1982; Deci and Ryan, 1985).

As in Plato's description of the different motivations that drive the human soul (see quotation above), recent research has also been highlighting the different possible motivations that drive man-not just love of money, but also moral values, and the love of praise.[1] Policies that aim to affect agents' behaviour may have very different outcomes, depending on which types of agents this policy addresses (see Tirole, 1994; Mishra and Samuel, 2016).

In this paper, we study how heterogeneity in individuals' motivations affects policies targeting corruption. Corruption is defined as bribe-taking behaviour while in public office. We assume that individuals not only have monetary incentives but also intrinsic motivation and reputation concerns to be honest.

[1] Mill (1909) firstly pointed the limitation of the economic approach in focusing exclusively on material incentives without concerns for every other human passion or motive. Fehr and Falk (2002) introduced individuals' motivation to reciprocate, the desire for social approval and the intrinsic enjoyment to work on interesting tasks to shape agents' behaviour. Caselli and Morelli (2003) developed a model of the elective office rewarded by 'ego rents', i.e., social status and power.

Simply put, reputation concerns stem from the inference drawn by the general public about an individual's motivations for performing an action. When monetary incentives increase, this inference is negatively affected and may lead to a crowding out of intrinsic motivation that is driven by image concerns. Titmuss, 1970 and Bénabou and Tirole, 2006 discuss the case of blood donation- when the monetary rewards for donating blood increase, those individuals who care about the perception of intentions may stop donating blood. This is referred to as the "overjustification effect" (Lepper and Greene, 1973): extrinsic incentives destroy the reputational value of good deeds.[2]

We build on Bénabou and Tirole, 2006 in formalising the "overjustification effect" and provide a simple theoretical foundation of the crowding out effect on reputational concerns for honesty that monetary incentives may induce. In the psychological literature, Lepper and Greene (1973) e.g. observed that "when an individual observes another person engaging in some activity, he infers that the other is intrinsically motivated to engage in that activity to the extent that he does not perceive salient, unambiguous, and sufficient extrinsic contingencies to which to attribute the other's behaviour". In line with this idea, we model collective reputation as the expected level of intrinsic motivation of honest agents. Professional reputation is high when honest behaviour is driven by intrinsic motivations and not by extrinsic incentives, like the fear of losing a well-paid job. Agents care about public perception of their degree of *intrinsic motivation to be honest relative to their own internal moral standard*, rather than by perception of honesty itself.[3] While it is hard to measure professional reputation defined in this way, public perception of the degree of intrinsic motivation for the job is plausibly positively correlated with occupational prestige: thus while firefighters figure high on occupational prestige, accountants, bankers, real estate agents have low prestige. We claim that this is at least partly because firefighters are paid less and thus it is more convincing that they are truly intrinsically motivated.[4]

[2]Rewards create doubt about the true motive for which prosocial behaviour is performed (Bénabou and Tirole, 2006).

[3]In the words of Adam Smith (1759) : "Nature, ... She taught him to feel pleasure in their favourable, and pain in their unfavourable regard. She rendered their approbation most flattering and most agreeable to him for its own sake, and their disapprobation most mortifying and most offensive."

[4]Further indirect evidence can be found in the survey made by Harris Interactive in 2014 looking at occupational prestige in the US the occupations with the highest prestige are: doctors (88% adults considering it to have either a great deal of prestige or prestige), military officers (78%), firefighters (76%), scientists (76%), and nurses (70%). At the other end of the spectrum, real estate broker/agent is the profession with the highest percentage of adults

Our first main result is to show conditions under which higher salaries may have a perverse effect on total corruption, even when there are no selection issues. The intuition behind this result is as follows. On the one hand, as is typical in standard models of corruption, a wage increase lowers "greedy" (i.e. extrinsically motivated) agents' incentives to accept bribes, because it raises the opportunity cost of being corrupt. This raises the level of honesty in the profession. On the other hand, outsiders' perception of the level of intrinsic motivation to be honest in the profession is *negatively* affected when salaries increase, because outsiders no longer attribute honesty to intrinsic motivations but rather to the fear of losing a well-paid job. This reduction of the professional reputation implies that reputation concerned agents have lower incentives to be honest as the opportunity cost of being corrupt goes down. The overall effect, therefore, depends on the relative weights that monetary and non-monetary motivations have in the reputation-concerned officers' utility function, as well as on their internal moral standards.

Our second main result is about selection: Suppose there is a private sector which only offers wage incentives but does not have any mission orientation so that intrinsic motivation plays no role. We then analyse the selection effects of a negative shock on private sector salaries. We show conditions under which selection of extrinsically motivated officers in the public sector causes a crowding out of intrinsic motivation due to reputation effects on reputation concerned agents. When the fraction of such agents is large enough, higher public sector wages may lead to higher corruption.

Overall, the paper highlights a mechanism which may explain why there is such conflicting evidence on the effect of higher wages to prevent bribe-taking in the public sector. It also suggests that rather than focusing on wages which can backfire, optimal policies to reduce corruption should take account of the competing motivations for public sector work and the importance of maintaining a high status in the public sector as opposed to private sector work. Even though each individual is motivated by monetary incentives, the net effect of non-monetary incentives is more straightforward than the use of monetary motivations.

The rest of the paper is organised as follows: Section 2 presents a brief literature review, Section 3 describes the model, Section 4 summarizes the main findings on how increasing salaries in the public sector affect total corruption.

considering it to have less prestige (73%), jointly with stockbrokers, bankers (62%), and accountants (60%).

Section 5 adds the private sector to the model. Finally Section 6 concludes.

2 Literature Review

Classic incentive literature argues that pure income maximisation is a core source of corruption (see e.g., Andvig and Moene, 1990; Acemoglu, 1995; Aidt, 2003; Sah, 2007) and the key to reducing corruption is to increase the minimum salary above what they can get elsewhere (Becker and Stigler, 1974; Chank and Moene, 1997), thereby increasing the opportunity cost of taking bribes. However, recent research suggests that wage increases may be an expensive and inefficient way to fight corruption (Rijckeghem and Weder, 2001, Bardhan, 2015) and may even encourage it (Foltz and Opoku-Agyemang, 2015). Indeed the recent literature on corruption highlights the role of non-monetary incentives, such as moral costs and social norms (see e.g., Polinsky and Shavell, 2000; Spichtig and Traxler, 2011). Our paper fits in with this latter approach where we take a more behavioural perspective on the motivations of public servants.

Our paper also ties into the large psychology literature on intrinsic motivation and crowding out. We use the psychological definition of intrinsic motivation, i.e., commencing an activity because it is absorbing and gratifying in itself, as opposed to initiating an activity for external prods, pressures, or rewards (Deci and Ryan, 1985). There is also substantial evidence from psychologists and sociologists that expected explicit rewards and punishments to perform a task sometimes may undermine a person's intrinsic motivation (e.g., Frey and Oberholzer-Gee, 1997; Deci, Koestner and Ryan, 1999). The crowding out effect has been used to explain the failure of incentive schemes in some principal-agent settings both theoretically (Bénabou and Tirole, 2003, Ellingsen and Johannesson, 2008) and empirically (Frey, 1997; Fehr and Gachter, 2001; Pokorny, 2008). Schulze and Frank (2003) show that the net effect of deterrence on overall corruption is *a priori* undetermined because deterrence increases the risk of being corrupt but erodes individual's intrinsic motivation for honesty. Our paper is similar in using crowding out effects but our focus is on image concerns or reputation concerns- we explicitly allow individuals to lose their desire to act pro-socially when their intentions are misattributed. Secondly, the presence of "greedy" agents in the pool, implies that higher monetary incentives crowd out intrinsic motivation. The negative externality imposed by extrinsically motivated agents in the pool causes crowding out in our model. This is similar to Tirole (1996) who argues that agents' incentives are not only

determined by their own reputation, but also by that of the groups they belong to. However, Tirole (1996) does not investigate the crowding out of honest behaviour due to the overjustification effect.

The paper is most closely related to Bénabou and Tirole (2006) (BT) in that agents value being perceived as having a high level of intrinsic motivation and gain negative reputation as being "greedy" or money-motivated and a higher incentive rate reduces the informativeness of actions about intrinsic motivation, while increasing it about monetary motivation. In other words, rewards create doubt about the true motive for prosocial behavior, and this "overjustification effect" can induce a crowding out of good behavior. In line with BT, we define the reputation payoff as others' belief as whether the action was taken due to pecuniary or intrinsic motivations (intention-based reputation payoffs). Building on BT, we assume that reputational benefit depends linearly on observers' posterior expectations of the agent's type and there is heterogeneity in individual's level of reputation concerns which creates noise in the "signal extraction" problem. However, our paper differs from BT in an important aspect. In our model, the individual payoff of prosocial behaviour does not depend on the public perception alone, but on the *difference* between the public perception and their own intrinsic motivation as well. As discussed earlier, the model recognises that it is not collective reputation for honest *behaviour* that matters but the motives that the public attributes to the behaviour. Moreover, it is collective reputation relative to the individual's internal moral standard that matters.

Finally, Machiavello's (2008) model assumes that workers are heterogeneous in terms of public sector motivation so that low public sector wage premia helps screen workers with intrinsic motivation, while high wage helps motivate workers to be honest. In his model, private sector wages are endogenous and depend on the quality of governance in the public sector. His model yields a non-monotonic relationship between wages and corruption: at low wages intrinsically motivated individuals are attracted to the public sector, as a result governance is good and private sector salaries are high- positive assortative matching takes place in equilibrium. If however, wages are too low then all agents are corrupt and private sector wages are also low. When wages increase in the public sector, it attracts some extrinsically motivated agents who cannot get high enough wages in the private sector, but will behave opportunistically in the public sector. This sustains the low wages in the private sector. At a basic level, the paper discusses how heterogeneous motivations imply that higher wages do not necessarily help to reduce corruption (increase the public sector motivation).

However the mechanism he highlights is different from ours: there are no image concerns, the behaviour of each type of agent depends only on wages and intrinsic motivation. Different from Machiavello (2008), our model would predict that public sector jobs can be differentiated according to whether they attract highly motivated individuals who also have image concerns or whether they attract individuals with low intrinsic motivation and high image concerns because the impact of higher wages varies according to this. While he does find a pattern - non monotonic relationship between wages and corruption, our main point is to explain why the empirical evidence does not find much evidence in either direction. We argue that this is because it will be hard to find systematic patterns as wages vary when individuals are motivated by image concerns about their intrinsic motivation.

Valasek (2016) follows Tirole's(1996) definition of collective reputation and shows that a non-monotonic wage path may be optimal to reform an institutional culture, attracting motivated agents. In Valasek's model, the level of collective reputation equals the fraction of honest agents, and therefore only depends on their behaviour and not on their motivation. Differently, in our model, collective reputation depends both on heterogeneity in agents' motives and behaviour.

3 The Model

3.1 Basic setup

Below we use the words individuals, agents and officers interchangeably. There is a mass 1 of individuals. We consider a standard model of corruption in which an officer, who receives a salary $w \geq 0$ for her job, decides whether or not to accept a bribe $B > 0$. If she accepts the bribe, then she is caught with a probability $q \in (0, \bar{q})$ with $\bar{q} < 1$. Let B denote the amount of exogenous bribe, q the probability of getting caught accepting a bribe and w the wage rate.

There are three types of officers $t \in \{G, R, S\}$: For simplicity of notation we refer to the three types as "greedy" officers (G), in proportion β, "reputation-concerned" officers (R) , in proportion α and "saints" (S), in proportion γ, where $\beta + \gamma + \alpha = 1$. All types care about the salary they get. Greedy officers care *only* about monetary incentives. Saints are highly intrinsically motivated, such that even with zero wage they would not accept a bribe. Reputation-concerned officers have a positive level of intrinsic motivation and are also motivated by

the desire for praise (reputation).

Let $v_t \geq 0$ denote type t officers' intrinsic valuation for being honest and $\theta_t \geq 0$ the weight on their concern for reputation. We assume $(v_G, \theta_G) = (0, 0)$, $(v_S, \theta_S) = (v^S, 0)$ with $v^S \geq (1-q)B - qw \equiv \bar{v}$, which guarantees that they never participate in corruption, and finally, $(v_R, \theta_R) = (v^R, \theta^R)$ with $v^R \in (0, v^S)$ and $\theta^R > 0$. Officers' type is private information, while the proportion of each type in the population, as the levels of v, and θ for each type, are common knowledge.

The utility function of an honest officer of type t is

$$U_t^{honest}(\delta) = w + v_t + \theta_t[qc(\delta) - v_t]. \tag{1}$$

The term $c(\delta)$ measures officers' collective reputation and depends on the behavior of each type and δ_t is the fraction of honest officers of type t and $\delta \equiv (\delta_G, \delta_R, \delta_S)$.

If a corrupt officer is charged, she gets zero utility; while if she is not, she gets

$$U_t^{corrupt}(\delta) = (1 - q)[w + B + \theta_t qc(\delta)].$$

We define the collective reputation $c(\delta)$ as the expected level of an honest officer's intrinsic motivation:[5]

$$c(\delta) \equiv \frac{\delta_R \alpha v_R + \delta_S \gamma v_S + \delta_G \beta v_G}{\delta_G \beta + \delta_R \alpha + \delta_S \gamma}.$$

As in Lepper and Greene (1973) and Bénabou and Tirole (2006), collective reputation is high when officers' honest behaviour is driven by intrinsic motivations and not by extrinsic incentives, like the fear of losing a well-paid job. We assume that the reputational payoff of an individual is the difference between the collective reputation discounted by q and her own intrinsic motivation.[6] By assumption, this is different from zero only for reputation-concerned officers.

Whether a certain level of praise is a reputational loss or gain for an $R-type$ officer depends on her own behaviour and the level of intrinsic motivation: an honest $R - type$ officer considers any level of praise lower than her intrinsic motivation as a reputational loss, $qc < v^R$, while for a corrupt officer any positive level of praise is a reputational gain. Thus an $R-type$ officer is motivated to be honest when the public rewards her for a higher level of honesty than her own

[5] Notice that, for any given equilibrium strategy profile, collective reputation is uniquely determined.

[6] When q is low (high), then the probability that an officer who is not caught is actually honest are low (high). Therefore, when the probability of detecting corruption is low in a country, the perceived collective reputation $qc(\delta)$ will be low.

intrinsic motivation to be honest (undeserved praise). On the other hand, the $R-type$ officer gets disutility from being honest when the public credits her with a lower level of intrinsic motivation than her actual level and is therefore encouraged to be corrupt (undeserved shame).[7] When honour or shame is felt it is only with regard to an internal moral compass in our model. This is important because individuals with high moral values may be systematically attracted to some types of jobs. There can be both crowding in and crowding out of intrinsic motivation in our model- the initial level of intrinsic motivation of R types in the population is key. This implies that it is in precisely those jobs which attract highly motivated individuals and have high occupational prestige (firefighters, nurses, scientists), that increases in salary have a deleterious effect.

3.2 Officers' behaviour

We want to analyse how a change in the salary offered to officers affect their propensity to accept a bribe, depending on their type. We restrict our analysis to symmetric equilibrium. From now on, let $\delta_t \in [0,1]$ be the probability that an officer of type t refuses the bribe. By assumption, $S-type$ officers never accept any bribe, and therefore in equilibrium $\delta_S = 1$. $G-type$ officers only care about monetary incentives and therefore their behaviour depends exclusively on the salary offered to them. A $G-type$ officer chooses to be honest if and only if

$$w \geq (1-q)(w+B), \tag{2}$$

defining a cutoff rule on the wage:

$$w \geq \frac{(1-q)}{q}B \equiv w_H^*. \tag{3}$$

Observe that as $q \to 0$, $w_H^* \to +\infty$, i.e, if the probability of being charged becomes negligible, the minimum salary necessary to keep a $G-type$ officer honest becomes arbitrarily large. The bribe is exchanged as long as the penalty, i.e., the probability of detection, is small enough (as shown in Basu et al., 2015). The following simple proposition directly follows from the above analysis.[8]

[7] If $v^R = 0$, i.e., R-type officials don't have intrinsic motivation, the analysis would correspond to the case in which they only care about the collective reputation. We discuss this case in the conclusion.

[8] The analysis could be extended to the case where $v^G > 0$. We prefer to keep the assumption that greedy officers are not intrinsically motivated to have a standard selfish type in our model. If $v^G > 0$ the main difference is that the threshold of the salary to keep greedy officers honest reduces, but our main results still hold.

Proposition 1 *A $G-type$ officer is honest if and only if $w \geq w_H^*$.*

$R-type$ officers' behaviour depends on their individual characteristics, the level of intrinsic motivation and their love of praise, as also on collective reputation. We need, then, to understand the relation between the salary paid to officers and their esteem in public eyes.

Consider first the case $w \geq w_H^*$, when, by Proposition 1, $G-type$ officers are honest. The following proposition highlights parameter values for which there exist equilibria where reputation-concerned officers may be corrupt even when material incentives are large enough to keep greedy officers honest. The first part highlights one of our main results on the overjustification effect: high wages and parameter values such that greedy agents are honest while reputation concerned officers are corrupt in equilibrium. We rule out situations where wages are so high that all agents are honest simply due to extrinsic motivations. Such salaries would be prohibitively expensive and moreover, when bribes are endogenous to salaries (e.g., see Mookherjee and Png (1995)) , it is not clear that such a salary exists.

The full characterization of the set of equilibria and related proofs are in the Appendix.

Proposition 2 *Suppose $w \geq w_H^*$. If $v^R \in (0, \frac{\gamma v^S}{1-\alpha})$,*[9] *then*
(1) $\delta_R = 1$ if $\theta^R \in (0, \theta_H^1]$ or $v^R \in (0, v_H^1]$;
(2) $\delta_R = 0$ if $\theta^R \geq \theta_H^0$ and $v^R \in (v_H^0, \frac{\gamma v^S}{1-\alpha})$;
(3) $\delta_R = \delta_H^{mix}$, where $\delta_H^{mix} \in (0,1)$, if $v^R \in (v_H^1, v_H^0)$, and $\theta^R > \theta_H^1$, or if $v^R \in (v_H^0, \frac{\gamma v^S}{1-\alpha})$ and $\theta^R \in (\theta_H^1, \theta_H^0)$.
If $v^R \in [\frac{\gamma v^S}{1-\alpha}, v^S)$, then
(4) $\delta_R = 1$ if $\theta^R \in (0, \theta_H^0]$;
(5) $\delta_R = 0$ if $\theta^R \geq \theta_H^1$.

Proof: *see the Appendix.*

The intuitive explanation of these results is the following: High salaries crowd out $R-type$ officers' individual intrinsic motivations, because an external observer cannot unambiguously attribute a high level of intrinsic motivation to honest behaviour. High salaries, therefore, negatively affect collective reputation. It follows that $R-type$ officers refuse the bribe if either (i) they have

[9] If $R-type$ officers' level of intrinsic motivation is high $v^R \in [\frac{\gamma v^S}{1-\alpha}, v^S)$, the collective reputation when $R-type$ officers choose to be honest is higher than the collective reputation when they choose to be corrupt; on the contrary, if their level of intrinsic motivation is low $v^R \in (0, \frac{\gamma v^S}{1-\alpha})$, the collective reputation is lower when $R-type$ officers choose to be honest.

a small love of praise, $\theta^R \leq \min\left\{\theta_H^0, \theta_H^1\right\}$ (these bounds, as also the bound v_H^1 below, are computed in the proof in the Appendix) and therefore they do not care about collective reputation even when it represents a reputation loss for them, or (ii) they have a low level of intrinsic motivation, $v^R \leq v_H^1$, so that they are similar to greedy officers; in this case collective reputation is still higher than their intrinsic motivation due to the presence of $S-type$ officers in the population. Therefore, there is a reputation gain in being honest- low intrinsic motivation agents benefit from undeserved praise.

On the other hand, if $R-type$ officers have a high reputation concern, $\theta^R \geq \max\left\{\theta_H^0, \theta_H^1\right\}$ and are sufficiently intrinsically motivated: $v^R \in [v_H^0, v^S)$, they will accept the bribe because the fact that an external observer assigns a high probability that an honest officer is the greedy type, crowds out their intrinsic motivation, making it costly for them to be honest: this is the effect of undeserved shame and the over justification effect.

Figure 1 shows these results. When the salary is high, $G-type$ officers choose to be honest, decreasing the level of collective reputation. An $R-type$ officer with a high level of intrinsic motivation feels a reputation loss and then has an incentive to be corrupt. An $R-type$ officer with a low level of intrinsic motivation feels a reputation gain and chooses to be honest no matter how much she cares about being praised. Image concerns make an officer with a high level of intrinsic motivation (high internal moral standard) more prone to be corrupt than an officer with the low level of intrinsic motivation.

The next proposition shows that this overjustification effect persists even when wage is lower so that greedy types are no longer honest. The first such case is when wages are still high enough to keep R-types honest when they do not care about reputation at all ($w \geq w_L^* \equiv \frac{(1-q)B}{q} - \frac{v^R}{q}$,). As wages decrease, the negative externality from $G-types$ dissappears. However, wages that guarantee that low intrinsic motivation agents are honest create a negative externality for R type individuals who have a high intrinsic motivation and high image concerns: the wages are still too high to rule out extrinsic motivations as the motivation to be honest. However, the range of parameters over which the effect of undeserved shame and overjustification occurs is smaller than before as wages are lower. This can be seen in Figures 2 and 3.

Proposition 3 *Suppose* $w \in [w_L^*, w_H^*)$*; then*
(1) $\delta_R = 1$ *if* $\theta^R \in (0, \theta_M^1]$ *or* $v^R \in (0, v_M^1]$*;*
(2) $\delta_R = 0$ *if* $\theta^R \geq \theta_M^0$ *and* $v^R \in (v_M^0, v^S)$*;*

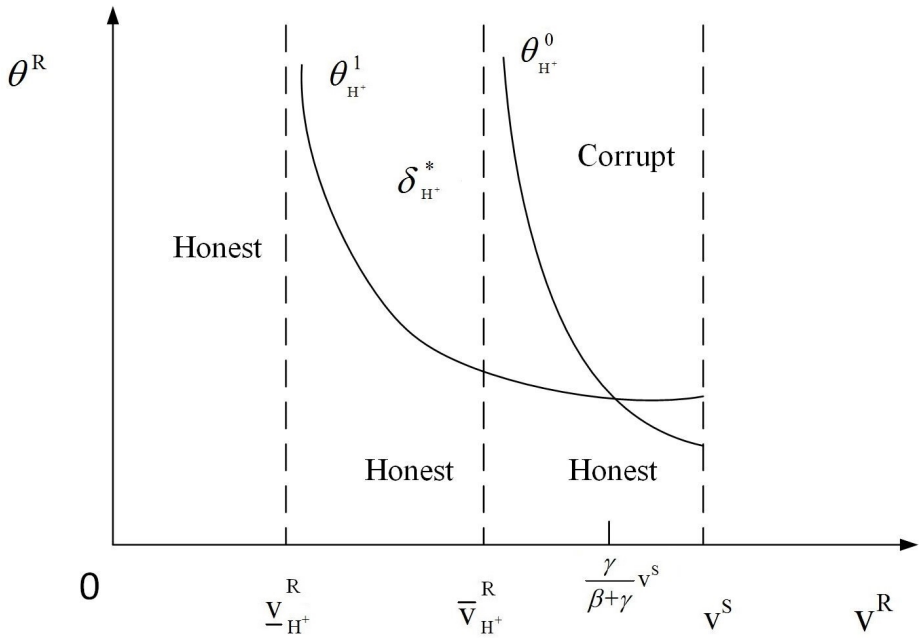

Figure 1: An illustration of $R-type$ officers' behavior when $G-type$ officers are honest ($w \geq w_H^*$).

(3) $\delta_R = \delta_M^{mix}$, *where* $\delta_M^{mix} \in (0,1)$, *if* $v^R \in (v_M^1, v_M^0)$ *and* $\theta^R > \theta_M^1$ *or if* $v^R \in (v_M^0, v^S)$ *and* $\theta^R \in (\theta_M^1, \theta_M^0)$.

Proof: see Appendix.

Finally, when $w < w_L^*$, then $R-type$ officers are honest only if they are less intrinsically motivated and have a strong love of praise: a large reputation gain is the only reward that can push them to be honest. As intrinsic motivation increases, if there is no concern for reputation, then an intrinsically motivated R type would be honest. However as soon as there is a positive concern for reputation, the collective reputation effect crowds out intrinsic motivation to be honest. The detailed description of all the equilibria in this range of parameters is contained in the following proposition.

Proposition 4 *Suppose* $w \in (0, w_L^*)$*; then*
(1) $\delta_R = 1$ *if* $\theta^R \geq \theta_L^1$ *and* $v^R \in (0, v_M^1)$*;*
(2) $\delta_R = 0$ *if* $\theta^R \in (0, \theta_L^0]$ *or* $v^R \in [v_M^0, v^S)$*;*
(3) $\delta_R = \delta_L^{mix}$, *where* $\delta_L^{mix} \in (0,1)$, *if* $v^R \in (0, v_M^1)$ *and* $\theta^R \in (\theta_L^0, \theta_L^1)$, *or if* $v^R \in (v_M^1, v_M^0)$ *and* $\theta^R > \theta_L^0$.

Proof: see Appendix.

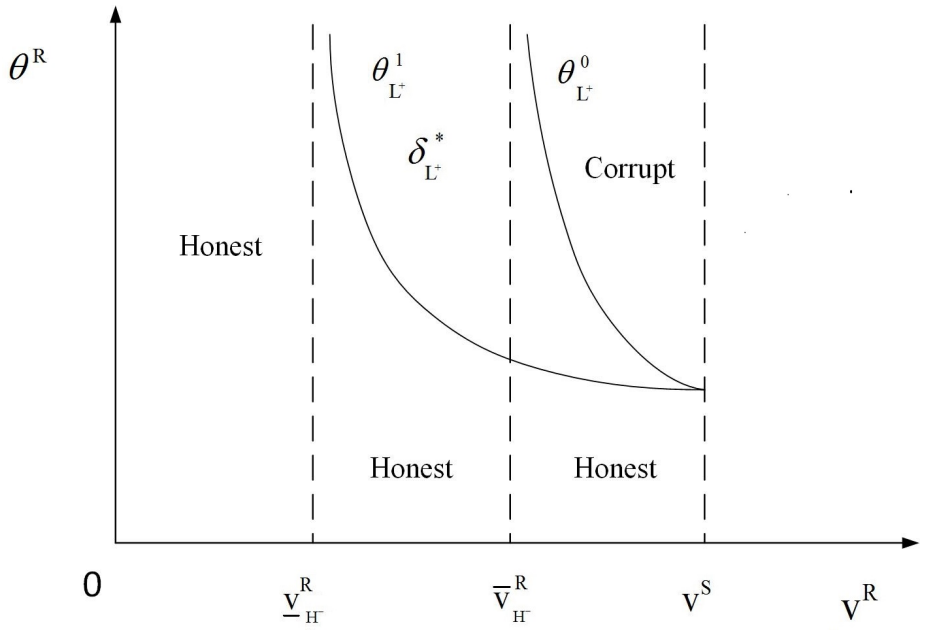

Figure 2: An illustration of $R-type$ officers' behavior when $G-type$ officers are dishonest and $w \in [w_L^*, w_H^*)$.

Overall, this section showed that when officers are motivated both intrinsically and by image/reputation concerns, then higher wages are not necessarily associated with higher probity in public life. If there is a large correlation between the level of intrinsic motivation and image concerns then the over-justification effect implies that even though greedy types are honest, many of the reputation concerned officers will not be as they suffer undeserved shame at being pooled with greedy types. The intuition behind this is that higher salaries lower the signal precision about intrinsic motivation (as in BT). Unlike BT, this effect persists even at lower salaries albeit the range of parameters for which it occurs is smaller. Unlike BT, jobs which attract individuals with low intrinsic motivation do not suffer the over-justification effect, because they benefit from undeserved praise from being pooled with saints. When wages are low, such individuals will be more motivated to work than those who are highly intrinsically motivated. Higher salaries help to make greedy types honest but have ambiguous effects on others depending on the level of intrinsic motivation and reputation concerns. The main driving force in our model is that image concerns are always relative to a personal moral standard: the overjustifica-

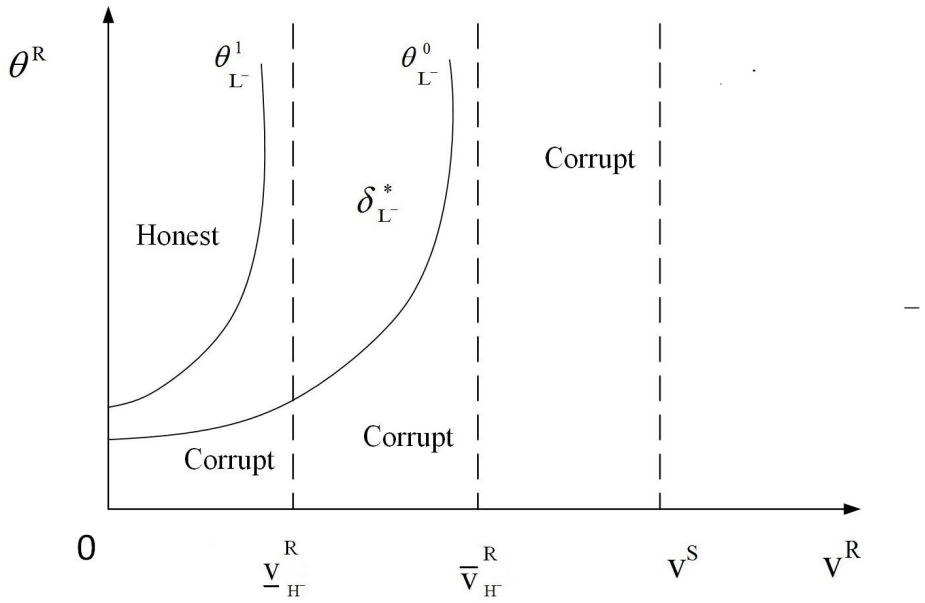

Figure 3: An illustration of $R-type$ officers' behavior when $G-type$ officers are dishonest and $w \in (0, w_L^*)$.

tion effect arises from being misunderstood in the public perception of one's intentions- sometimes it has a positive effect and sometimes a negative one.

4 The impact of increasing officers' salary on the level of corruption

So far we have described the various equilibria that exist at different intervals of the salary. We are interested in the overall effect of increasing officers' salary on the level of corruption but this is not easily predicted. While $G-type$ officers' propensity to accept a bribe unambiguously decreases when the salary increases, the way that $R-type$ officers react to a salary increase depends on the levels of their intrinsic motivation and the love of praise. Thus, whether officers will be more or less prone to accept a bribe after a pay rise depends on their type, and the overall impact will depend on the proportion of each type among officers' population. The following two propositions illustrate the different effects that increasing salary may have on the level of corruption, measured as the fraction of officers who accept a bribe. Increasing salaries weakly reduces $G-type$ offi-

cers' propensity of being corrupt. Therefore, anytime a pay rise lowers $R-type$ officers' propensity to corruption, it reduces the level of corruption. This happens when R-type officers have sufficiently low moral standards or sufficiently low weight on reputation concerns.

Proposition 5 *Suppose that either*

(i) $v^R \in (0, v_H^1]$ *and* $\theta^R < \theta_L^1$*; or*

(ii) $v^R \in (\max\left\{v_M^0, \frac{\gamma v^s}{1-\alpha}\right\}, v^S)$ *and* $\theta^R < \min\left\{\theta_M^0, \theta_H^0\right\}$*. If the salary increases from any* $w \leq w_L^*$ *to any* $w \geq w_L^*$,[10] *then the level of corruption decreases.*

Proof: see Appendix.

In case (i) $R-type$ officers have a low level of intrinsic motivation as well as low reputation concerns. A pay rise from $w \in (0, w_L^*)$ to $w \in [w_L^*, w_H^*)$ does not affect $G-type$ officers' behaviour because they are corrupt when the salary is below the threshold w_H^*. However, the rising salary modifies the behaviour of $R-type$ officers who, after the pay rise, stop accepting bribes, as described in Proposition 3. When $R-type$ officers have a low level of intrinsic motivation, an increase in salary works in the same direction as an increase in reputation. This is because $R-type$ officers with low moral standards are motivated by a reputational gain (undeserved praise). A further pay rise turns the $G-type$ officers honest and therefore reduces the propensity of being corrupt of every officer to 0. In case (ii), $R-type$ officers are sufficiently intrinsically motivated and care very little about reputation. Not surprisingly, given our previous discussion, this is the case in which the crowding out effect of a pay rise is very limited and therefore the direct positive effect of a pay rise prevails.

Next, we present our main result showing how an increase in salary that makes greedy agents honest may lead to higher corruption among the R-types, because it crowds out their motivation by reducing the collective reputation of their profession and thereby reducing the precision of the signal.

The following propositions highlight the trade-off that a policy to fight corruption may have to confront: a pay rise may reduce $G-type$ officers' propensity to be corrupt, while increasing $R-type$ officers' propensity of being corrupt because it crowds out their motivation by reducing the collective reputation of their profession.

[10] We focus on comparison of two public salary intervals but not on incremental increases in salary levels because most policies of salary increases to fight corruption are not marginal increases. Empirical research (e.g., Rijckeghem and Weder, 2001; Foltz and Opoku-Agyemang, 2015) also focuses on the effectiveness of considerably increases in salary to reduce corruption rather than incremental increases.

Proposition 6 *Suppose that either*
(i) $v^R \in [v_H^1, v_M^1]$, $\theta^R > \theta_H^1$, *and* $\alpha > \frac{\beta}{1-\delta_H^{mix}}$; *or*
(ii) $v^R \in [v_H^0, \min\{v_M^0, \frac{\gamma v^S}{1-\alpha}\}]$, $\theta^R > \max\{\theta_M^1, \theta_H^0\}$, *and* $\alpha > \frac{\beta}{\delta_M^{mix}}$.
In both cases, if the salary increases from the interval $w \in [w_L^*, w_H^*)$ *to* $w \geq w_H^*$, *the level of corruption increases.*

Proof: see Appendix.

In case (i) above, an increase in salary that reduces officers' collective reputation turns $R-type$ officers, who have an intermediate level of intrinsic motivation and strong reputational concerns, from being corrupt with probability zero to being corrupt with positive probability δ_H^{mix}. In case (ii), an increase in salary that reduces officers' collective reputation turns $R-type$ officers from being honest with positive probability, δ_M^{mix} to being honest with probability zero. In both cases, a pay rise increases $R-type$ officers' propensity of being corrupt. If there are sufficiently many agents of this type, even if a pay-rise reduces $G-type$ officers' propensity of being corrupt, the overall effect of a pay rise is counter-productive, and the level of corruption increases.

The above propositions highlight how difficult is to predict the effect of a pay rise on the level of corruption. Still, our analysis offers some guidelines to policymakers, about the cases when a salary increase may have a paradoxical effect on the level of corruption.

Remark 1 *he higher is the fraction of officers who are highly reputation-concerned and have high moral standards, the more likely a pay rise may have the perverse effect of increasing corruption among officers.*

5 The choice of working in the public sector: Selection

Until now we have assumed that all types of agents work in the public sector. We now focus on selection into the public sector as a function of relative salary. Our goal in this section is to show that selection issues may be an additional reason which causes crowding out of motivation of R types. When the salary in the public sector increases relative to the private sector, it will attract greedy types who will be corrupt unless the public sector salary is high enough. This lowers the reputation payoff for R types who are highly intrinsically motivated and the overjustification effect kicks in. We see that starting from a low wage, low corruption scenario when the public sector wage relative to the private sector

wage goes up then corruption increases due to selection effects. Policies to prevent this from happening are either to increase salaries in the public sector until greedy types are no longer corrupt, however, this will cause R types to become corrupt and this policy will cause a big burden on the public exchequer. Alternatively, salaries can be reduced such that greedy types do not enter the public sector so that both criteria can be fulfilled at the same time. We show that this conclusion depends crucially on the composition of R types in the pool- if they do not care about reputation or have low intrinsic motivation then an increase in salary may be the optimal policy to reduce corruption. Either way, even if corruption reduces due to an increase in salary, it may crowd out the intrinsic motivation of the R types. We show this in the cases below with some simplifying assumptions.

There are two sectors in the economy. A public sector job provides intrinsic motivation, as before but a job in the private sector does not provide any intrinsic motivation[11]. Thus, the only motivation to work in the private sector is the monetary compensation. Each worker decides whether to work in the public sector, of size $z < 1$, or in the private sector of size $1 - z$. We consider a large stationary economy ($-\infty < t < +\infty$) in which at each time t the public sector re-recruits z officers. If there are more applicants to the public sector, candidates are randomly selected and those who are not selected end up working in the private sector.

We assume that the large majority of agents are greedy so β tends to one, $\gamma < z$, that is the number of incorruptible agents (saints) is not enough for the needs of the public sector, but $\alpha + \gamma > z$, which implies that the number of reputation-concerned agents and saints is larger than the size of the public sector. In the private sector, firms pay an exogenous competitive salary w^P. The salary in the public sector is chosen in order to minimise (i) the cost of hiring z officers (ii) the amount of corruption. A benevolent planner chooses the optimal policy that meets these two criteria, according to a social welfare function. Let w^G denote the salary paid in the public sector (government).

This simple model allows an analysis of the agents' choice of which sector they want to be employed in, to design the optimal wage offered in the public sector. Furthermore, it makes simple predictions of the effects that a negative shock on the private sector will have on the level of corruption in the public sector. We first consider an ideal situation in which a planner manages to

[11] See i.e., Wilson (1989), Tirole (1994), Sheehan (1996) and Besley and Ghatak (2005).

attract only motivated agents in the public sector and takes advantage of this matching to have low salaries and no corruption.

A greedy agent who works in the public sector behaves honestly only if $w^G \geq \frac{(1-q)}{q}B$. Let w^H be the minimum salary that keeps reputation-concerned officers honest when saints and reputation-concerned officers apply to the public sector and greedy agents do not apply to the public sector:

$$w^H + v^R + \theta^R \left[\frac{q(\gamma v^S + \alpha v^R)}{\alpha + \gamma} - v^R \right] = (1-q) \left[w^H + B + \frac{q\theta^R(\gamma v^S + \alpha v^R)}{\alpha + \gamma} \right]. \tag{4}$$

Rearranging (4), we get

$$w^H = \frac{(1-q)}{q}B - \frac{q\theta^R(\gamma v^S + \alpha v^R)}{\alpha + \gamma} - \frac{v^R(1-\theta^R)}{q}. \tag{5}$$

The salary w^H is lower than the salary that is needed to keep a greedy officer honest ($w^H \leq \frac{(1-q)}{q}B$). If (i) $v^R \leq \frac{q^2\gamma v^S}{\alpha+\gamma-q^2\alpha} \equiv v^{CR}$,[12] or (ii) $v^R > v^{CR}$ and $\theta^R \leq \frac{v^R}{v^R - \frac{q^2(\gamma v^S + \alpha v^R)}{\gamma+\alpha}} \equiv \theta^{CR}$ then R-type officers are honest. In the first case, reputation-concerned officers have a low level of intrinsic motivation and they get a reputation gain from the collective reputation of the public sector. In the second case, they are motivated by a high level of intrinsic motivation and do not have a strong love of praise. In the case that only saints and reputation-concerned officers work in the public sector, these officers have an additional motive to behave honestly rather than only being motivated by monetary incentives, which allows a planner to offer low salaries. Let

$$w_1^P \equiv \frac{(1-q)}{q}B - \frac{v^R(1-q)(1-\theta^R)}{q}$$

and

$$w_0^P \equiv w_1^P - \frac{q\theta^R(1-q)(\gamma v^S + \alpha v^R)}{\alpha + \gamma}$$

Proposition 7 *If $w_0^P \leq w^P \leq w_1^P$, $v^S \geq \max\{\bar{v}^{HS}, \hat{v}^S(w^p)\}$ and either of the following conditions holds: (i) $v^R \leq v^{CR}$, or; (ii) $v^R > v^{CR}$ and $\theta^R \leq \theta^{CR}$, the optimal compensation policy to minimise corruption in the public sector is to offer a salary equal to w^H identified by (5). In equilibrium all greedy workers work in the private sector and motivated workers work in the public sector. The level of corruption is zero.*

Proof: *see the Appendix.*

[12] This threshold is the same as $\underline{v}_{H^-}^R$ identified by (41) in Appendix to prove Proposition 3.

Consider now the consequences of an economic shock that produces a fall in the competitive salary offered in the private sector. Namely, suppose that the salary in the private sector reduces to $\tilde{w}^p < (1-q)(w^H + B)$: now working in the public sector is attractive for greedy officers.

If the salary in the public sector does not change, then corruption arises directly because greedy agents apply for a job in the public sector and choose to be corrupt. The fraction of corrupt officers is equal to β, which by assumption tends to one. Since corruption is widespread, a benevolent planner has two policies to fight corruption. The first policy aims to increase the salary in the public sector to raise the opportunity costs of corruption for greedy officers. The minimum salary that induces greedy officers to be honest is $w^G = \frac{1-q}{q}B$. In this case, the collective reputation of a job in the public sector tends to zero because citizens assign probability near to one that an honest officer is a greedy type. A reputation-concerned officer prefers to accept a bribe if

$$\frac{(1-q)}{q}B + v^R(1-\theta^R) \leq (1-q)[\frac{(1-q)}{q}B + B]$$

which clearly holds if and only if $\theta^R \geq 1$.[13]

In this case, the overall level of corruption tends to zero (because by assumption β tends to one), but reputation-concerned officers' motivations are crowded out.

Lemma 1 *Suppose that $\theta^R \geq 1$ and (i) the salary in the public sector is equal to $w^G = \frac{1-q}{q}B$ and (ii) the salary in the private sector is $w^P \leq \frac{(1-q)}{q}B + v^R(1-q\theta^R)$. Hence, all types of agents apply to the public sector. Greedy officers and saints behave honestly, while reputation-concerned officers are corrupt.*

A policy that implies an increase in the remuneration of public officers during a recession may be hard to implement. An alternative policy is to reduce the salary in the public sector in order to discourage greedy agents to apply to the public sector. Let $\tilde{w}^p$ denote the salary paid in the private sector satisfying condition (ii) in Lemma 1. If the planner decides to lower the salary in the public sector, then reputation-concerned officers will be corrupt and the level of corruption will be $\frac{\alpha}{\alpha+\gamma}$. This discussion can be summarised by the following remark.

Remark 2 *Economic recession attracts non-motivated agents into the public sector. A benevolent planner can either react by increasing salaries in the public sector that incentivise non-motivated agents to be honest. If most of the*

[13]Note that $\theta^{CR} > 1$. Therefore, this condition $\theta^R \geq 1$ is compatible with the condition previously imposed that $\theta^R \leq \theta^{CR}$.

agents are greedy, the planner restores a negligible level of corruption, even if motivations for reputation-concerned agents are crowded out. Alternatively, the planner can lower the salary in the public sector, to discourage greedy agents to apply, but it induces a positive level of corruption among reputation-concerned officers.

A shock on the private sector has a similar effect as a pay rise in the public sector because it makes it more attractive to work in the public sector for greedy officers. For simplicity, we start from a situation in which only motivated agents work in the public sector and the (technical) assumptions in Proposition 7 guarantee that this is the case. If it is possible to only attract motivated workers, the optimal salary is the minimum that keeps motivated (reputation-concerned) officers honest. There is a shock which causes the salary in the private sector to drop and therefore greedy officers are attracted by the public sector. Policies that the planner may implement to fight corruption can make motivated officers dishonest.

This simple extension of our model with two sectors shows, as pointed out in the previous section, that there may exist situations in which greedy officers are honest while reputation-concerned officers are corrupt. More importantly, it points out that the remuneration policy in the public sector is not only relevant to design incentives to fight corruption, but also to determine the composition of the workforce. Corruption is not only a problem of moral hazard but also of adverse selection, and this latter consideration has not received the same attention than the former one.

6 Conclusion

Why does corruption take place in the public sector? We introduce a model based on the key premise that public officers' behaviour is driven by a mix of motivations: monetary incentives, intrinsic value for honesty and the love of praise. This mix varies across individuals, and, importantly, the distributions of types among the population affects individual behaviour through its effect on collective reputation. A pay rise that lowers greedy officers' propensity to be corrupt may have an opposite effect on reputation-concerned officers, because it crowds out officers' collective reputation by questioning their true motive to be honest.

It is important to point out that it is the presence of image concerns about *intentions* and the internal moral standards that are important for the result- in

an alternative interpretation, suppose prestige is defined differently and is higher when fewer officers are caught being corrupt.[14] Then, there is no trade off- higher salaries imply lower corruption for everyone absent any selection issues. Assume now that prestige is defined as the collective reputation about the level of intrinsic motivation to be honest, but that all R types are positively affected by reputation regardless of initial level of intrinsic motivation (the analysis would correspond to the case in which $v^R = 0$)- higher salaries will still lead to crowding out as collective reputation goes down due to pooling with G types. But of course sufficiently high salaries imply lower corruption.

We summarise the main message of this paper in the following way: policies that aim to undermine the benefits of corruption have to be coupled with those that make working for the public good more prestigious, as Paul Collier lucidly state in a recent policy paper (2016): *"Until well into the 19th century, the British public sector was very corrupt. [...]. By the late 19th century, the British Civil Service had become honest and competent. This transformation was largely fortuitous rather than the result of a properly thought-through strategy. But its success reveals the key components of how change can be brought about. [....]In Britain, two key things – closing off the major opportunities for corruption and making working for the public good more prestigious and satisfying than abusing office for private gain – happened together. These two approaches are jointly critical in breaking cultures of corruption."*

7 Appendix

Proof of Proposition 2:

Given that $R-type$ officers refuse to accept a bribe in equilibrium. In this case, the level of collective reputation is equal to

$$\frac{\alpha v^R + \gamma v^S}{\alpha + \beta + \gamma} = \alpha v^R + \gamma v^S \equiv c_H^1 \tag{6}$$

This is a Nash equilibrium if the following inequality holds:

$$w + v^R + \theta^R(qc_H^1 - v^R) \geq (1-q)(w + B + q\theta^R c_H^1) \tag{7}$$

Substituting (6) into (7), leads to the following expression

$$\theta^R[q^2(\alpha v^R + \gamma v^S) - v^R] \geq (1-q)B - qw - v^R. \tag{8}$$

[14] This definition of course has some problems about conflating lower enforcement with lower perceived corruption, but let us ignore that for the argument.

Since by hypothesis $w \geq w_H^*$, from (3), we get

$$(1-q)B - qw - v^R < 0. \tag{9}$$

The right side of equation (8) is negative and, since $\theta^R > 0$ by assumption, it follows that (7) is valid whenever

$$q^2(\alpha v^R + \gamma v^S) - v^R \geq 0, \tag{10}$$

Condition (10) defines a cutoff v_H^1,

$$0 < v^R \leq \frac{q^2\gamma v^S}{1-q^2\alpha} \equiv v_H^1, \tag{11}$$

such that for any level of intrinsic motivation $v^R \in (0, v_H^1]$, an $R-type$ officer chooses to be honest with probability one no matter how much she cares about being praised. If $v \in (v_H^1, v^S)$, we get

$$q^2(\alpha v^R + \gamma v^S) - v^R < 0, \tag{12}$$

then an $R-type$ officer chooses to be honest if and only if $\theta^R \in (0, \theta_H^1]$, where

$$\theta_H^1 \equiv \frac{-(1-q)B + qw + v^R}{v^R - q^2(\alpha v^R + \gamma v^S)}. \tag{13}$$

An $R-type$ public officer whose intrinsic motivation is $v^R \in (v_H^1, v^S)$ chooses to be honest with probability one if and only if $\theta^R \in (0, \theta_H^1]$.

When $w \geq w_H^*$, an $R-type$ officer chooses to be corrupt if and only if

$$w + v^R + \theta^R(qc_H^0 - v^R) \leq (1-q)(w + B + q\theta^R c_H^0) \tag{14}$$

where $c_H^0 \equiv \frac{\gamma v^S}{\beta+\gamma}$ is collective reputation when saints and greedy officers are honest; substituting c_H^0 into (14), leads to the following expression

$$\theta^R\left(\frac{q^2\gamma v^S}{\beta+\gamma} - v^R\right) \leq (1-q)B - qw - v^R, \tag{15}$$

from (9) and given that $\theta^R > 0$, (15) is satisfied when

$$\frac{q^2\gamma v^S}{\beta+\gamma} - v^R < 0. \tag{16}$$

We obtain a cutoff value

$$v^R > \frac{q^2\gamma v^S}{\beta+\gamma} \equiv v_H^0 \tag{17}$$

Therefore, if $v^R \in (v_H^0, v^S)$, an $R-type$ officer chooses to be corrupt if and only if $\theta^R \in [\theta_H^0, +\infty)$, where

$$\theta_H^0 \equiv \frac{-(1-q)B + qw + v^R}{v^R - \frac{q^2\gamma v^S}{\beta+\gamma}}. \tag{18}$$

When $w \geq w_H^*$, an $R-type$ officer whose intrinsic motivation $v^R \in (0, \frac{\gamma v^S}{1-\alpha})$ is indifferent between being honest and corrupt if and only if there is a $\delta_H^{mix} \in (0,1)$ such that

$$w + v^R + \theta^R(qc_H^{mix} - v^R) = (1-q)(w + B + q\theta^R c_H^{mix}), \tag{19}$$

where

$$c_H^{mix} \equiv \frac{\delta_H^{mix}\alpha v^R + \gamma v^S}{\delta_H^{mix}\alpha + \beta + \gamma}. \tag{20}$$

Substituting (20) into (19), we obtain

$$\delta_H^{mix} = \frac{(\beta+\gamma)[(1-q)B - qw - (1-\theta^R)v^R] - q^2\theta^R\gamma v^S}{\alpha\{q^2\theta^R v^R - [(1-q)B - qw - (1-\theta^R)v^R]\}}. \tag{21}$$

For any $v^R \in (v_H^0, \frac{\gamma v^S}{1-\alpha})$, if

$$(1-q)B - qw - (1-\theta^R)v^R < q^2\theta^R v^R, \tag{22}$$

then

$$(\beta+\gamma)[(1-q)B - qw - (1-\theta^R)v^R] - q^2\theta^R\gamma v^S < (\beta+\gamma)q^2\theta^R v^R - q^2\theta^R\gamma v^S. \tag{23}$$

By the assumption $0 < v^R < \frac{\gamma v^S}{\beta+\gamma}$, we obtain

$$\theta^R q^2[(\beta+\gamma)v^R - \gamma v^S] < 0, \tag{24}$$

and by (22) and (23), $\delta_H^{mix} < 0$, which implies that the solution to (19) does not belong to $(0,1)$. $\delta_H^{mix} > 0$ holding requires

$$q^2\theta^R v^R - [(1-q)B - qw - (1-\theta^R)v^R] < 0, \tag{25}$$

which implies

$$\theta^R > \frac{v^R + qw - (1-q)B}{v^R(1-q^2)} \equiv \hat{\theta}_H. \tag{26}$$

For all $\theta^R > \hat{\theta}_H$, $\delta_H^{mix} > 0$ if and only if

$$(\beta+\gamma)[(1-q)B - qw - (1-\theta^R)v^R] - q^2\theta^R\gamma v^S < 0. \tag{27}$$

If $v^R \in (v_H^1, v_H^0)$, we get $v^R - \frac{q^2\gamma v^S}{\beta+\gamma} < 0$, and (27) holds for all $\theta^R > \hat{\theta}_H$.

For all $\theta^R > \hat{\theta}_H$, suppose $v^R \in (v_H^1, \frac{\gamma v^S}{1-\alpha})$, by (25), $\delta_{H^+}^* < 1$ if and only if

$$(\beta+\gamma)[(1-q)B-qw-(1-\theta^R)v^R]-q^2\theta^R\gamma v^S > \alpha\{q^2\theta^R v^R-[(1-q)B-qw-(1-\theta^R)v^R]\}, \tag{28}$$

which equals to

$$\theta^R[q^2(\alpha v^R + \gamma v^S) - v^R] < (1-q)B - qw - (1-\theta^R)v^R < 0. \tag{29}$$

(29) defines a cutoff value

$$\theta^R > \frac{-(1-q)B + qw + v^R}{v^R - q^2(\alpha v^R + \gamma v^S)} \equiv \theta_H^1 > \hat{\theta}_H. \tag{30}$$

Therefore, if $v^R \in (v_H^1, v_H^0)$, there exists a $\delta_H^{mix} \in (0,1)$ if and only if $\theta^R > \theta_H^1$.

Suppose $v^R \in [v_H^0, \frac{\gamma v^S}{1-\alpha})$, we get $v^R - \frac{q^2\gamma v^S}{\beta+\gamma} > 0$, (27) defines a cutoff value

$$\theta^R < \frac{-(1-q)B + qw + v^R}{v^R - \frac{q^2\gamma v^S}{\beta+\gamma}} \equiv \theta_H^0. \tag{31}$$

Suppose $\theta^R > \hat{\theta}_H$ and $v^R \in (v_H^0, \frac{\gamma v^S}{1-\alpha}]$, (29) holds if and only if

$$(\beta+\gamma)[(1-q)B-qw-(1-\theta^R)v^R]-q^2\theta^R\gamma v^S > \alpha\{q^2\theta^R v^R-[(1-q)B-qw-(1-\theta^R)v^R]\}, \tag{32}$$

which defines a cutoff value

$$\theta^R > \frac{-(1-q)B + qw + v^R}{v^R - q^2(\alpha v^R + \gamma v^S)} \equiv \theta_H^1 > \hat{\theta}_H. \tag{33}$$

To conclude, if $v^R \in (v_H^0, \frac{\gamma v^S}{1-\alpha})$, an $R-type$ officer chooses to be honest with probability $\delta_H^{mix} \in (0,1)$ if and only if $\theta^R \in (\theta_H^1, \theta_H^0)$.

Furthermore, we analyze the case in which $v^R \in [\frac{\gamma v^S}{1-\alpha}, v^S)$. Suppose $v^R \in [\frac{\gamma v^S}{1-\alpha}, v^S)$, we have $c_H^1 \geq c_H^0$. Then, we get $0 < \theta_H^0 \leq \theta_H^1$. The above analysis implies that in this case, an $R-type$ officer chooses to be honest if and only if

$$0 < \theta^R \leq \theta_H^0, \tag{34}$$

and chooses to be dishonest if and only if

$$\theta^R \geq \theta_H^1. \tag{35}$$

When $\theta^R \in (\theta_H^0, \theta_H^1)$, there is no equilibrium. □

Proof of Proposition 3:

Define $w_L^* \equiv \frac{(1-q)B}{q} - \frac{v^R}{q}$, when $w \in [w_L^*, w_H^*)$, an $R-type$ officer chooses to be honest ($\delta_R = 1$) if and only if

$$w + v^R + \theta^R(qc_M^1 - v^R) \geq (1-q)(w + B + q\theta^R c_M^1), \tag{36}$$

where

$$c_M^1 = \frac{\alpha v^R + \gamma v^S}{\alpha + \gamma}. \tag{37}$$

Substituting (37) into (36), leads to the following expression

$$\theta^R[\frac{q^2(\alpha v^R + \gamma v^S)}{\alpha + \gamma} - v^R] \geq (1-q)B - qw - v^R. \tag{38}$$

Since by hypothesis $w \geq w_L^*$, from the definition of w_L^*, we get

$$(1-q)B - qw - v^R \leq 0. \tag{39}$$

The right side of equation (38) is non-positive and, since $\theta^R > 0$ by assumption, it follows that (36) is valid whenever

$$\frac{q^2(\alpha v^R + \gamma v^S)}{\alpha + \gamma} - v^R \geq 0, \tag{40}$$

Condition (40) defines a cutoff v_M^1,

$$0 < v^R \leq \frac{q^2\gamma v^S}{1 - q^2\alpha - \beta} \equiv v_M^1, \tag{41}$$

such that for any level of intrinsic motivation $v^R \in (0, v_M^1]$, an $R-type$ officer chooses to be honest with probability one no matter how much she cares about being praised.

If $v^R \in (v_M^1, v^S]$, we get

$$\frac{q^2(\alpha v^R + \gamma v^S)}{\alpha + \gamma} - v^R < 0, \tag{42}$$

then an $R-type$ officer chooses to be honest if and only if $\theta^R \in (0, \theta_M^1]$, where

$$\theta_M^1 \equiv \frac{-(1-q)B + qw + v^R}{v^R - \frac{q^2(\alpha v^R + \gamma v^S)}{\alpha + \gamma}}. \tag{43}$$

An $R-type$ officer whose intrinsic motivation is $v^R \in (v_M^1, v^S)$ chooses to be honest if and only if $\theta^R \in (0, \theta_M^1]$.

When $w \in [w_L^*, w_H^*)$, an $R-type$ chooses to be dishonest ($\delta_R = 0$) if and only if

$$w + v^R + \theta^R(qc_M^0 - v^R) \leq (1-q)(w + B + q\theta^R c_M^0), \tag{44}$$

where

$$c_M^0 = v^S, \tag{45}$$

is collective reputation when only saints are honest; substituting (45) into (44), leads to the following expression which leads to

$$\theta^R(q^2 v^S - v^R) \leq (1-q)B - qw - v^R, \tag{46}$$

from (39) and given that $\theta^R > 0$, (46) is satisfied when

$$v^R \geq q^2 v^S \equiv v_M^0. \tag{47}$$

Therefore, if $v^R \in (v_M^0, v^S)$, an $R-type$ officer chooses to be corrupt if and only if $\theta^R \in [\theta_M^0, +\infty)$, where

$$\theta_M^0 \equiv \frac{-(1-q)B + qw + v^R}{v^R - q^2 v^S}. \tag{48}$$

Therefore, an $R-type$ officer whose intrinsic motivation $v^R \in (v_M^0, v^S)$ chooses to be dishonest if and only if $\theta^R \in [\theta_M^0 + \infty)$ where $\theta_M^0 > \theta_M^1 > 0$.

When $w \in [w_L^*, w_H^*)$, an $R-type$ public officer is indifferent between honest and dishonest if and only if there is a $\delta_M^{mix} \in (0,1)$ such that

$$w + v^R + \theta^R(qc_M^{mix} - v^R) = (1-q)(w + B + q\theta^R c_M^{mix}), \tag{49}$$

where

$$c_M^{mix} \equiv \frac{\delta_M^{mix}\alpha v^R + \gamma v^S}{\delta_M^{mix}\alpha + \gamma}. \tag{50}$$

Substitute (50) into (49), get

$$\delta_M^{mix} = \frac{\gamma\{q^2\theta^R v^S - [(1-q)B - qw - (1-\theta^R)v^R]\}}{\alpha\{[(1-q)B - qw - (1-\theta^R)v^R] - q^2\theta^R v^R\}}. \tag{51}$$

If

$$(1-q)B - qw - (1-\theta^R)v^R < q^2\theta^R v^R, \tag{52}$$

by $0 < v^R < v^S$, we get

$$(1-q)B - qw - (1-\theta^R)v^R < q^2\theta^R v^R < q^2\theta^R v^S. \tag{53}$$

$\delta_M^{mix} < 0$ violates the assumption that $\delta_M^{mix} \in (0,1)$. Therefore, $\delta_M^{mix} > 0$ if and only if

$$q^2\theta^R v^R < (1-q)B - qw - (1-\theta^R)v^R < q^2\theta^R v^S. \tag{54}$$

Define $\hat{\theta}_M \equiv \frac{v^R+qw-(1-q)B}{(1-q^2)v^R} > 0$, if $v^R \in (v_M^1, v_M^0)$, $\delta_M^{mix} > 0$ holds for all $\theta^R > \hat{\theta}_M$. If $v^R \in (v_M^0, v^S)$, (54) defines a cutoff

$$\hat{\theta}_M < \theta^R < \frac{v^R + qw - (1-q)B}{v^R - q^2 v^S} \equiv \theta_M^0. \tag{55}$$

For all $\theta^R > \hat{\theta}_M$, $\delta_M^{mix} < 1$ if and only if

$$\gamma\{q^2\theta^R v^S - [(1-q)B - qw - (1-\theta^R)v^R]\} < \alpha\{[(1-q)B - qw - (1-\theta^R)v^R] - q^2\theta^R v^R\}, \tag{56}$$

so we get

$$\theta^R[\frac{q^2(\alpha v^R + \gamma v^S)}{\alpha+\gamma} - v^R] < (1-q)B - qw - v^R. \tag{57}$$

If $v^R \in (v_M^1, v^S)$, (57) defines a cutoff

$$\theta^R > \frac{-(1-q)B + qw + v^R}{v^R - \frac{q^2(\alpha v^R + \gamma v^S)}{\alpha+\gamma}} \equiv \theta_M^1, \tag{58}$$

and $\theta_M^1 > \hat{\theta}_M$.

Therefore, suppose $v^R \in (v_M^1, v_M^0)$, there exists a $\delta_M^{mix} \in (0,1)$ if and only if $\theta^R > \theta_M^1$. If $v^R \in (v_M^0, v^S)$, $\delta_M^{mix} \in (0,1)$ exists if and only if $\theta^R \in (\theta_M^1, \theta_M^0)$. □

Proof of Proposition 4:

When $w \in (0, w_L^*)$, an $R-type$ officer chooses to be dishonest ($\delta_R = 0$) if and only if

$$w + v^R + \theta^R(qc_M^0 - v^R) \leq (1-q)(w + B + q\theta^R c_M^0), \tag{59}$$

where

$$c_M^0 = v^S. \tag{60}$$

Substituting (60) into (59), leads to the following expression

$$\theta^R(q^2 v^S - v^R) \leq (1-q)B - qw - v^R. \tag{61}$$

By the hypothesis $0 < w < w_L^*$, we get

$$(1-q)B - qw - v^R > 0. \tag{62}$$

The right side of equation (61) is positive and since $\theta^R > 0$ by the assumption, it follows that (59) is valid whenever

$$v^R \geq q^2 v^S \equiv v_M^0 \tag{63}$$

For any level of intrinsic motivation $v^R \in [v_M^0, v^S)$, an $R-type$ officer chooses to be dishonest with probability one no matter how much she cares about being praised.

If $v^R \in (0, v_M^0)$, we get

$$q^2 v^S - v^R > 0, \tag{64}$$

then an $R-type$ officer chooses to be dishonest if and only if $\theta^R \in (0, \theta_L^0]$ where

$$\theta_L^0 \equiv \frac{(1-q)B - qw - v^R}{q^2 v^S - v^R}. \tag{65}$$

An $R-type$ officer whose intrinsic motivation $v^R \in (0, v_M^0)$ chooses to be dishonest if and only if $\theta^R \in (0, \theta_L^0]$.

When $w \in (0, w_L^*)$, an $R-type$ officer chooses to be honest ($\delta_R = 1$) if and only if

$$w + v^R + \theta^R(qc_M^1 - v^R) \geq (1-q)(w + B + q\theta^R c_M^1), \tag{66}$$

where

$$c_M^1 = \frac{\alpha v^R + \gamma v^S}{\alpha + \gamma}. \tag{67}$$

Substituting (67) into (66), we get

$$\theta^R \left[\frac{q^2(\alpha v^R + \gamma v^S)}{\alpha + \gamma} - v^R \right] \geq (1-q)B - qw - v^R. \tag{68}$$

From (62) and given that $\theta^R > 0$, (68) is satisfied when

$$v^R < \frac{q^2 \gamma v^S}{1 - \beta - q^2 \alpha} \equiv v_M^1. \tag{69}$$

Therefore, if $v^R \in (0, v_M^1)$, an $R-type$ officer chooses to be honest if and if $\theta^R \in [\theta_L^1, \infty)$ where

$$\theta_L^1 \equiv \frac{(1-q)B - qw - v^R}{\frac{q^2(\alpha v^R + \gamma v^S)}{\alpha + \gamma} - v^R}. \tag{70}$$

When $w \in (0, w_L^*)$, an $R-type$ officer is indifferent between honest and dishonest if and only if there is a $\delta_L^{mix} \in (0,1)$ such that

$$w + v^R + \theta^R(qc_M^{mix} - v^R) = (1-q)(w + B + q\theta^R c_M^{mix}), \tag{71}$$

where

$$c_M^{mix} \equiv \frac{\delta_L^{mix}\alpha v^R + \gamma v^S}{\delta_L^{mix}\alpha + \gamma}. \tag{72}$$

Substituting (72) into (71), we get

$$\delta_L^{mix} = \frac{\gamma\{q^2\theta v^S - [(1-q)B - qw - (1-\theta^R)v^R]]\}}{\alpha\{[(1-q)B - qw - (1-\theta^R)v^R] - q^2\theta^R v^R\}}. \tag{73}$$

For any $w \in (0, w_L^*)$, from (62), we get $(1-q)B - qw - v^R + \theta^R(1-q^2)v^R > 0$ holds for all $\theta^R > 0$. Therefore, $\delta_L^{mix} > 0$ if and only if

$$(1-q)B - qw - (1-\theta^R)v^R < q^2\theta^R v^S. \tag{74}$$

If $v^R \in (0, v_M^0)$, then (74) defines a cutoff

$$\theta^R > \frac{(1-q)B - qw - v^R}{q^2 v^S - v^R} \equiv \theta_L^0. \tag{75}$$

And $\delta_L^{mix} < 1$ if and only if

$$\gamma\{q^2\theta^R v^S - [(1-q)B - qw - (1-\theta^R)v^R]\} < \alpha\{[(1-q)B - qw - (1-\theta^R)v^R] - q^2\theta^R v^R\}, \tag{76}$$

which leads to

$$\theta[\frac{q^2(\alpha v^R + \gamma v^S)}{\alpha + \gamma} - v^R] < (1-q)B - qw - v^R. \tag{77}$$

If $v^R \in (v_M^1, v_M^0)$, (77) holds for any $\theta^R > 0$. If $v^R \in (0, v_M^1)$, then (77) defines a cutoff

$$\theta^R < \frac{(1-q)B - qw + v^R}{\frac{q^2(\alpha v^R + \gamma v^S)}{\alpha+\gamma} - v^R} \equiv \theta_L^1. \tag{78}$$

When $w \in (0, w_L^*)$, suppose $v^R \in (0, v_M^0)$, there exists a $\delta_L^{mix} \in (0,1)$ if and only if $\theta^R \in (\theta_L^0, \theta_L^1)$. If $v^R \in (v_M^1, v_M^0)$, there exists a $\delta_L^{mix} \in (0,1)$ if and only if $\theta^R > \theta_L^0$. □

Proof of Proposition 5: (i) Suppose $v^R \in (0, v_H^1]$, $\theta^R < \theta_L^1$, and $w \in (0, w_L^*)$, $R-type$ officers choose to be corrupt with positive probability. By Proposition 1, $G-type$ officers are corrupt. When the salary increases to $w \in [w_L^*, w_H^*)$, $G-type$ officers are still corrupt. By Proposition 3, $R-type$ officers with $v^R \in (0, v_H^1]$ and $\theta^R < \theta_L^1$ choose to be corrupt with probability zero. Therefore, the level of corruption decreases to β. When the salary further increases to $w \geq w_{H*}$, $G-type$ officers become honest. $R-type$ officers with

$v^R \in (0, v_H^1]$ and $\theta^R < \theta_L^1$ still choose to be honest with probability 1. The level of corruption is zero.

(ii) By Proposition 4, suppose $v^R \in (\max\left\{v_M^0, \frac{\gamma v^s}{1-\alpha}\right\}, v^S)$, and $\theta^R < \min\left\{\theta_M^0, \theta_H^0\right\}$, when the salary is $w \in (0, w_L^*)$, $R-type$ officers choose to be dishonest with probability 1. By Proposition 1, $G-type$ officers are corrupt. Therefore, the proportion of corrupt public officer is $\alpha+\beta$. When the salary increases to $w \in [w_L^*, w_H^*)$, $G-type$ officers are still corrupt. $R-type$ officers with $v^R \in (\max\left\{v_M^0, \frac{\gamma v^s}{1-\alpha}\right\}, v^S)$ and love of praise $\theta^R < \min\left\{\theta_M^0, \theta_H^0\right\}$ choose to be honest with positive probability. The level of corruption decreases. When the salary further increases to $w \geq w_H*$, $G-type$ officers become honest. $R-type$ officers with $v^R \in (\max\left\{v_M^0, \frac{\gamma v^s}{1-\alpha}\right\}, v^S)$ and $\theta^R < \min\left\{\theta_M^0, \theta_H^0\right\}$ chooses to be honest with probability 1 so that all three types are honest. The level of corruption further decreases. □

Proof of Proposition 6:

(i) When the salary is $w \in [w_L^*, w_H^*)$, $G-type$ officers are dishonest. $R-type$ officers whose intrinsic motivation $v^R \in [v_M^1, v_H^1]$ and love of praise $\theta^R > \theta_H^1$ choose to be honest with probability one. The proportion of corrupt public officers is β. When the salary increases to $w \geq w_H*$, $G-type$ officers choose to be honest. $R-type$ officers whose intrinsic motivation $v^R \in [v_M^1, v_H^1]$ and love of praise $\theta^R > \theta_H^1$ choose to be corrupt with a positive probability δ_H^*. The proportion of corrupt public officer is $(1-\delta_H^{mix})\alpha$. Furthermore, if $\beta < (1-\delta_H^{mix})\alpha$, the level of corruption increases.

(ii) When the salary is $w \in [w_L^*, w_H^*)$, $G-type$ officers are dishonest. $R-type$ officers whose intrinsic motivation $v^R \in [v_H^0, \min\{v_M^0, \frac{\gamma v^S}{1-\alpha}\}]$ and $\theta^R > \max\{\theta_M^1, \theta_H^0\}$ choose to be honest with probability δ_M^{mix}. The level of corruption is $\beta+\alpha(1-\delta_M^{mix})$. When the salary increases to $w \geq w_H*$, $G-type$ officers choose to be honest. $R-type$ officers with $v^R \in [v_H^0, \min\{v_M^0, \frac{\gamma v^S}{1-\alpha}\}]$ and $\theta^R > \max\{\theta_M^1, \theta_H^0\}$ choose to be honest with probability zero. The level of corruption is α. If $\alpha > \frac{\beta}{\delta_M^{mix}}$, the level of corruption increases. □

Proof of Proposition 7: A greedy officer does not apply to the public sector if

$$w^p \geq \max\left\{w^G, (1-q)(w^G+B)\right\}. \tag{79}$$

A saint applies to the public sector if

$$w^G + v^S \geq w^P, \tag{80}$$

and chooses to be honest if

$$w^G + v^S \geq (1-q)(w^G + B). \tag{81}$$

An honest reputation-concerned agent applies to the public sector if

$$w^G + v^R + \theta^R \left[\frac{q(\gamma v^S + \alpha v^R)}{\alpha + \gamma} - v^R \right] \geq w^P. \tag{82}$$

If the salary offered in the public sector is $w^H \leq \frac{(1-q)}{q} B$, greedy officers prefer to work in the private sector if $w^P \geq (1-q)(w^H + B)$, or

$$w^P \geq \frac{(1-q)}{q} B - \frac{q\theta^R(1-q)(\gamma v^S + \alpha v^R)}{\alpha + \gamma} - \frac{v^R(1-\theta^R)(1-q)}{q} \equiv w_0^P.$$

Suppose $w^P \geq (1-q)(w^H + B)$, we have $w^P \geq w^H$. Conditions (80) and (81) require that, given w^P, there exists a $\hat{v}^S(w^P)$ such that for all $v^S \geq \hat{v}^S(w^P)$, $w^H + v^S - w^P \geq 0$ and a $\bar{v}^{HS}$ such that for all $v^S \geq \bar{v}^{HS}$, we have $qw^H + v^S - (1-q)B \geq 0$.

The fourth condition holds if

$$w^P \leq \frac{(1-q)}{q} B - \frac{v^R(1-q)(1-\theta^R)}{q} \equiv w_1^P$$

When the salary in the public sector is fixed at $w = w^H$, all officers behave honestly and collective reputation of the public sector is equal to $\frac{\gamma v^S + \alpha v^R}{\alpha + \gamma}$. At a lower salary reputation-concerned officers will be corrupt, therefore w^H is the minimum salary that allows a corruption free public sector. □

References

[1] Acemoglu, D. (1995) Reward Structures and the Allocation of Talent, *European Economic Review* 39 (1), 17-33.

[2] Aidt, T. (2003) Economic Analysis of Corruption: a Survey, *The Economics Journal* 113 (491), 632-652.

[3] Alesina, A., Danninger, S., Rostagno, M. (1999) Redistribution Through Public Employment: the Case of Italy. NBER WP, 7387.

[4] Amir, R., and C. Burr (2015) Corruption and Socially Optimal Entry, *Journal of Public Economics* 123, 30-41.

[5] Andvig, J., and K. Moene (1990) How Corruption May Corrupt, *Journal of Economic Behavior & Organization* 13 (1), 63-76.

[6] Bardhan, P. (2015). Corruption and Development Policy (Drawing Upon the Recent Indian Debate), *Journal of Public Economic Theory* 17(4), 472-479.

[7] Bardhan, P. (1997). Corruption and Development: A Review of Issues, *Journal of Economic Literature* XXXV (5), 893-921.

[8] Basu, K., K. Basu, and T. Cordella (2016). Asymmetric Punishment as an Instrument of Corruption Control, *Journal of Public Economic Theory* 18 (6), 831-856.

[9] Becker, G., and G. Stigler (1974) Law Enforcement, Malfeasance, and Compensation of Enforcers *The Journal of Legal Studies* 3, 1-18.

[10] Bénabou, R., and J. Tirole (2003) Intrinsic and Extrinsic Motivation, *Review of Economic Studies* 70, 489-520.

[11] Bénabou, R., and J. Tirole (2006) Incentives and Prosocial Behavior. *American Economic Review* 96 (5), 1652-1678.

[12] Cadot, O. (1987) Corruption As A Gamble, *Journal of Public Economics* 33, 223-244.

[13] Caselli, F., and M. Morelli (2003) Bad Politicians, *Journal of Public Economics* 88, 759-782.

[14] Chand, S., and K. Moene (1999) Controlling Fiscal Corruption, *World Development* 27 (7), 1129-1140.

[15] Collier, P. (2016) "How to change cultures of corruption"in *Against Corruption: a collection of essays*, Prime Minister's Office, 10 Downing Street, London, UK.

[16] Deci, E., R. Koestner, and R. Ryan (1999) A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic rewards on Intrinsic Motivation, *Psychological Bulletin* 125 (6), 627-668.

[17] Deci, E., and R. Ryan (1971) Effects of Externally Mediated Rewards on Intrinsic Motivation, *Journal of Personality and Social Psychology* 18, 105-115.

[18] Deci, E., and R. Ryan (1985). *Intrinsic Motivation and Self-determination in Human Behavior*, New York: Plenum.

[19] Dhillon, A. (2015) Introduction to the Special Issue on Governance and Political Economy, *Journal of Public Economic Theory* 17, 4, 469-471.

[20] Ellingsen, T., and M. Johannesson (2008) Pride and Prejudice: The Human Side of Incentive Theory, *American Economic Review* 98 (3), 990-1008.

[21] Fehr, E., and A. Falk (2002). Psychological Foundations of Incentives, *European Economic Review* 46, 4-5, 687-724.

[22] Fehr, E., and S. Gächter (2001) Do Incentive Contracts Crowd Out Voluntary Cooperation? Discussion Paper, No. 3017, CEPR.

[23] Frey, B. (1997) Not Just for the Money: An Economic Theory of Motivation, Edward Elgar Pub.

[24] Frey, B., and F. Oberholzer-Gee (1997) The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out, *American Economics Review* 87 (4), 746-755.

[25] George, A., and W. Dickens (1982) The Economic Consequences of Cognitive Dissonance, *American Economic Review* 72 (3), 307-319.

[26] Mahmood, M. (2005) Corruption in Civil Administration: Causes and Cures, *Humanomics* 21(3/4), 62-84.

[27] Macchiavello, R. (2008) Public Sector Motivation and Development Failures, *Journal of Development Economics* 86 (1), 201-213.

[28] Mill, J. (1909). *Principles of Political Economy with some of their Applications to Social Philosophy*, London: Longmans, Green and Co.

[29] Mishra, A., and A., Samuel (2017) Preemptive Bribery and Incomplete Information: Does Prior Knowledge Matter? *Journal of Public Economic Theory* 19, 658-673.

[30] Plato, 1985. *The Republic*, Translated by Jowett, B., First Avenue Editions.

[31] Pokorny, K. (2008) Pay-But Do Not Pay Too Much: An Experimental Study on the Impact of Incentives, *Journal of Economic Behavior & Organization* 66 (2), 251-264.

[32] Polinsky, M., and S. Shavell (2000) The fairness of sanctions: some implications for optimal enforcement policy, *American Law and Economics Review* 2 (2), 1223-237.

[33] Sah, R. (2007) Corruption Across Countries and Regions: Some Consequences of Local Osmosis, *Journal of Economic Dynamics and Control* 31 (8), 2573-2598.

[34] Seabright, P. (2004). Continuous Preferences Can Cause Discontinuous Choices: An Application to the Impact of Incentives on Altruism, Working Paper, No 257, Institut d'Ëconomie Industrielle.

[35] Schulze, G., and B. Frank (2003) Deterrence Versus Intrinsic Motivation: Experimental evidence on the Determinants of Corruptibility, *Economics of Governance* 4, 143-160.

[36] Smith, A. (1759) *The Theory of Moral Sentiments*, Amherst, New York: Prometheus Books.

[37] Swamy, A., S. Knack, and Y. Lee, and O. Azfar (2001) Gender and Corruption *Journal of Development Economics* 64 (1), 25-55.

[38] Tirole, J. (1994) The Internal Organization of Government, *Oxford Economic Papers* 46 (1), 1-29.

[39] Tirole, J. (1996) A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality, *Review of Economic Studies* 63 (1), 1-22.

[40] Titmuss, R. (1970) *The Gift Relationship: From Human Blood to Social Policy*, London: Allen and Unwin.

[41] Traxler, C., and M. Spichtig (2011). Social Norms and the Indirect Evolution of Conditional Cooperation, *Journal of Economics* 102 (3), 237-262.

[42] Treisman, D. (2000). The Causes of Corruption: A Cross-National Study, *Journal of Public Economics* 76 (3), 399-457.

[43] Valasek, J. (2016) Reforming an Institutional Culture: A Model of Motivated Agents and Collective Reputation, Working Paper No. 5599, CESifo.

[44] Van Rijckeghem, C., and B. Weder (2001) Bureaucratic Corruption and the Rate of Temptation: Do Wages in the Civil Service Affect Corruption, and by How Much? *Journal of Development Economics* 65 (2), 307-331.